

Sensors **2013**, *13*, 16894–16914; doi:10.3390/s131216894

OPEN ACCESS

sensors

ISSN 1424-8220

www.mdpi.com/journal/sensors

Article

A New Blondin System for Surveying and Photogrammetry

Federico Cuesta ^{1,*}, Francisco M. Lopez-Rodriguez ¹ and Antonio Esteban ²

¹ Escuela Técnica Superior de Ingeniería, University of Seville, Camino de los Descubrimientos s/n, Seville E-41092, Spain; E-Mail: francisco.m.lopez.rodriguez@juntadeandalucia.es

² Sacyr S.A.U., Av. Luis Montoto 107, Seville E-41007, Spain; E-Mail: aesteban@sacyr.com

* Author to whom correspondence should be addressed; E-Mail: fede@cartuja.us.es;
Tel.: +34-954-487-356; Fax: +34-944-487-340.

Received: 16 October 2013; in revised form: 18 November 2013 / Accepted: 3 December 2013 / Published: 6 December 2013

Abstract: The main objective of the system presented in this paper is to provide surveyors and engineers with a new photogrammetry device that can be easily integrated with surveying total stations and a global navigation satellite system (GNSS) infrastructure at a construction site, taking advantage of their accuracy and overcoming limitations of aerial vehicles with respect to weight, autonomy and skilled operator requirements in aerial photogrammetry. The system moves between two mounting points, in a blondin ropeway configuration, at the construction site, taking pictures and recording the data of the position and the orientation along the cable path. A cascaded extended Kalman filter is used to integrate measurements from the on-board inertial measurement unit (IMU), a GPS and a GNSS. Experimental results taken in a construction site show the system performance, including the validation of the position estimation, with a robotic surveying total station, or the creation of a digital surface model (DSM), using the emergent structure from motion (SfM) techniques and open software. The georeferencing of the DSM is performed based on estimated camera position or using ground control points (GCPs).

Keywords: sensors integration; surveying; construction site; cascaded extended Kalman filter (EKF); structure from motion (SfM); dense point cloud; digital surface model (DSM)

1. Introduction

Advances in sensors technologies have had a significant impact during the last decades on the ground-level tools used by surveyors and construction engineers in their daily tasks, including surveying, layout and grade management of a construction project [1]. Thus, nowadays, devices with millimetric or centimetric accuracy, such as digital theodolites, total stations, robotic total stations, terrestrial laser scanners (TLS) or real-time kinematics global navigation satellite systems (RTK-GNSS), are common tools on any construction site. Recently, Gikas [2] showed the benefits and limitations of using robotic total stations and TLS technology, to collect high-fidelity data to support tunnel construction in two construction sites in Greece. Park *et al.* [3] developed a monitoring system for complex building construction using a laser displacement sensor (LDS) and wireless sensors, whose accuracy and reliability were verified with a total station. Siu *et al.* [4] used two robotic total stations and three cameras to provide a solution to monitor the dimensional changes of a temporary facility in the construction field. In [5], El-Omari and Moselhi make use of a 3D laser scanner and photogrammetry, in an integrated way, to collect data from construction sites and to compare actual work to planned work in a building under construction. Fregonese *et al.* [6] analyzed the use of TLS and total stations for the surveying and monitoring of civil construction and historical buildings.

Improvements become evident not only for ground-level devices, but also for airborne-based tools, especially in aerial photogrammetry, mainly due to advances in digital images, CMOS camera sensors, MEMS inertial measurement units (IMU), integrated IMU-GPS devices, laser sensors, *etc.* Many applications have been reported about using airborne devices [7], including aerial mapping with handheld sensors [8], land surveying of construction sites [9], automated determination of 3D city models from airborne acquired point cloud data [10], cadastral applications [11], forest inventory [12], photogrammetric surveys in archaeological areas [13] or the generation of georeferenced digital terrain models based on structure from motion approaches [14].

Several types of flying objects, such as conventional planes, helicopters [8], kites [15] or a wide variety of unmanned aircraft systems (UAS) [7,16,17], had been used. At this point, UAS deserve special attention (see [16,17] and the references therein for a complete classification and considerations of use). Among them, micro UAVs, which require no takeoff or landing run, have experienced a significant spread in the last decade, mainly boosted by lightweight, low-cost MEMS-based IMU, GPS, cameras or lasers. The counter-part of these low-cost devices is a reduction of accuracy and performance, which can be mitigated by using sensor fusion techniques for pose estimation (both position and orientation) [12,18,19]. On the other hand, these vehicles usually have a reduced payload (lower than 3 kg) and limited autonomy. UAVs with accurate on-board devices, augmented payload and autonomy can also be found, but at a significantly higher cost [16]. Another key point is that these vehicles usually require a skilled operator, either to remote control or as the safety pilot in the case of UAVs with autopilot systems (in some cases, the autopilot can only be activated when the vehicle is in the air, but take-off and landing have to be operated manually) [16].

This paper presents a new system designed to take advantage of ground device accuracy (facilitating the integration with surveying total stations, RTK-GNSS and the construction site infrastructure) and, at the same time, overcoming the limitations of aerial vehicles with respect to weight, autonomy and

the need for skilled operators. Thus, the system integrates an accurate RTK-GNSS, commonly used at construction sites, a prism for interfacing with a robotic total station, a high resolution camera and a medium-cost IMU into a device, weighing around 12 kg in total. It moves between two mounting points, in a blondin ropeway configuration, at the construction site, taking pictures and recording the position and orientation data along the cable, as shown in Figure 1. Using two mounting points can be a disadvantage with respect to the coverage area, but this could be solved by using a Skycam configuration. The prototype has been developed in a research project funded by the Spanish construction company, Sacyr SAU.

Figure 1. System at the construction site.



This paper is organized as follows: Section 2 describes the sensors on board the system, as well as their integration using a cascaded extended Kalman filter. Section 3 shows some experimental results at a construction site, including the validation of position estimation, with a robotic surveying total station,

or the creation of a digital surface model (DSM), using structure from motion (SfM) techniques and open software. Georeferencing of the DSM is performed based on the estimated camera position or by means of ground control points (GCPs). Section 4 ends with the conclusions.

2. System Description and Sensors Integration

As mentioned above, the main objective of the system presented in this paper is to provide surveyors and engineers with a new photogrammetry device that can be easily integrated with the surveying total stations and GNSS infrastructure at the construction site. Thus, by including in the system an omnidirectional prism and a commonly used centimetric RTK-GNSS, it is possible to take advantage of the high precision provided by the total stations and available GNSS corrections. Moreover, conventional equipment facilitates the usage and integration of surveying operators (see Figure 2).

Figure 2. Surveyor at work configuring a conventional Topcon GR-3 real-time kinematics global navigation satellite system (RTK-GNSS) unplugged from the system (**left**) and recording the system position by using a robotic total station (**right**).

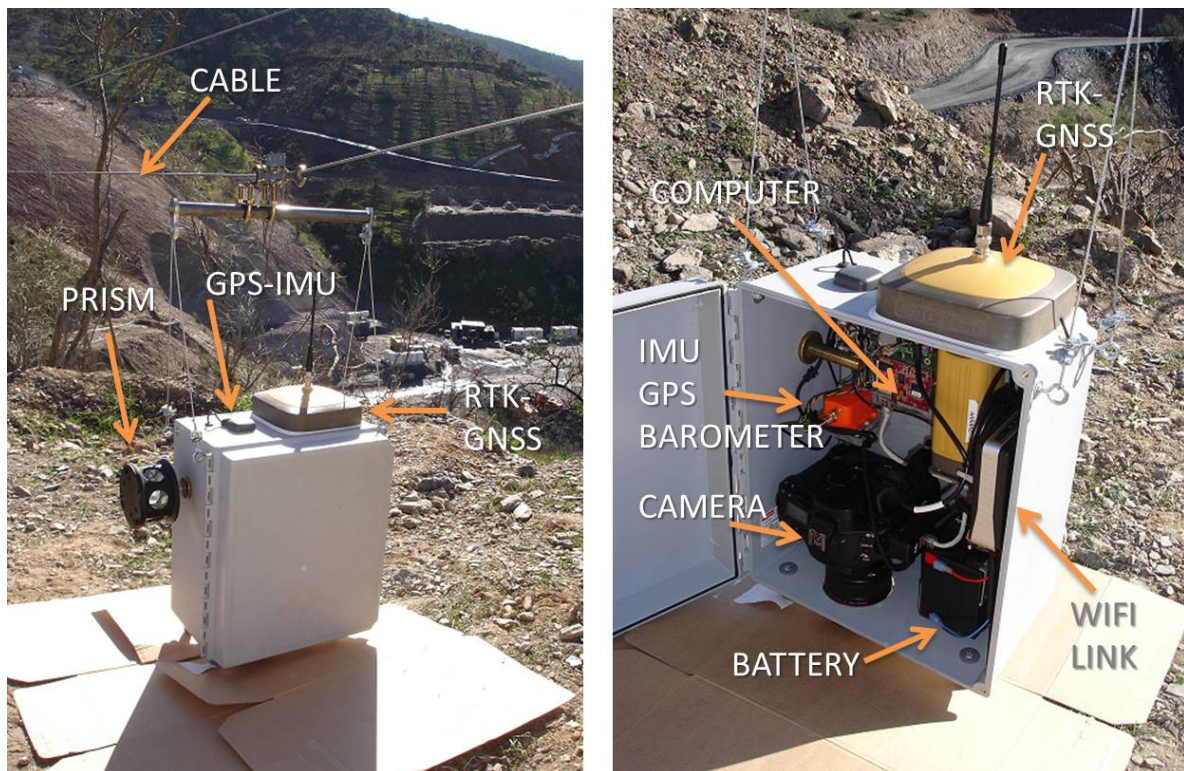


From a functional point of view, the system is fixed to a cable in such a way that it can move between two mounting points, in a blondin ropeway configuration, taking pictures and recording data of the position and orientation along the cable path. The system is connected to a remote PC, via a directional WiFi antenna (see Figure 2 (right)), for telemetry data logging and camera recording.

The system has a wide variety of sensors (see Figure 3), including 3D-accelerometers, 3D-rate gyroscopes, 3D-magnetometers, barometer, thermometer and GPS, all of them integrated into a commercial medium-low cost IMU-GPS unit (Xsens MTI-G), as well as a high resolution RTK-GNSS (Topcon GR-3), a high resolution professional photo camera (Canon EOS-1Ds Mark III), a 360 degree prism and a Wi-Fi link. An embedded computer collects and integrates data from all the sensor devices, and it also connects with the remote PC. Sensors are enclosed in an IP66compliant box, for protection against dust and water. The total weight of the system (batteries included) is around 12 kg, with an autonomy of several hours.

Concerning the IMU-GPS, it is a MEMS-based unit (with solid state capacitive readout accelerometers and rate of turn sensors and a silicon bulk micro-machined pressure sensor), which is able to provide heading and attitude with a dynamic accuracy of three and 1° RMS, respectively (the static accuracy is lower than one and 0.5°). The angular resolution is 0.05° . On the other hand, the GPS position accuracy is 2.0 m CEP (Differential GPS mode). The 3D magnetometer and barometer are used in an internal real-time Kalman filter as aiding sensors to directly observe magnetic north (improving heading estimation) and to account for varying pressure (which affects altitude estimation).

Figure 3. Sensors on board the system.



On the other hand, Topcon GR-3 is an advanced GPS+ (triple constellation GPS, GLONASS, GALILEO receiver), specially designed for surveying and engineering applications. In real-time RTK mode, the accuracy is $10 \text{ mm} + 1 \text{ ppm}$ (horizontal) and $15 \text{ mm} + 1 \text{ ppm}$ (vertical). It should be pointed out that the static accuracy can be improved with post-processing up to $+3.0 \text{ mm} + 0.5 \text{ ppm}$ (horizontal) and $+5.0 \text{ mm} + 0.5 \text{ ppm}$ (vertical). In differential mode, the accuracy is lower than 0.5 m in real time (and higher than 0.25 m with post-processing). Obviously, it is a high cost unit.

The on-board camera (Canon EOS-1Ds Mark III with a Canon EF 14 mm f/2.8L II USM Ultra-Wide Angle Lens) is a medium-high cost, professional digital camera, with a 21.1 Megapixel image sensor. Images can be obtained in raw format with a resolution of $5,616 \times 3,744$ pixels. An effective sensor size is $36 \times 24 \text{ mm}$, and the lens focal length, 14 mm.

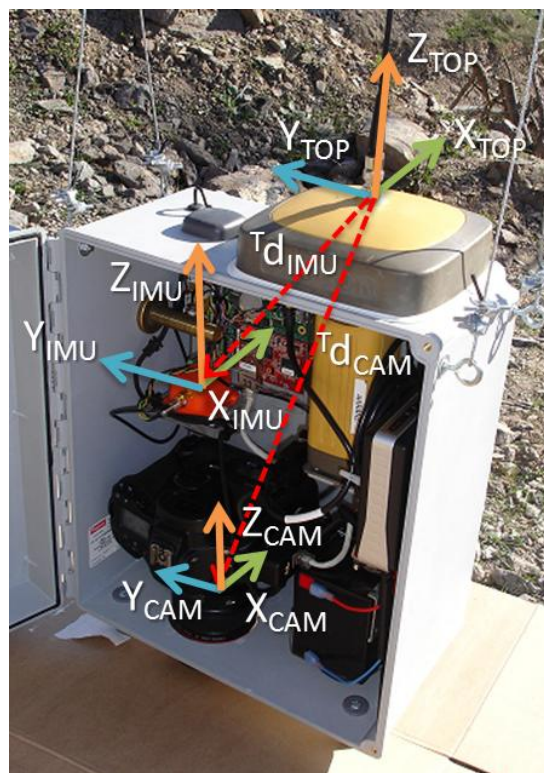
2.1. Sensors Integration and Extended Kalman Filter

Sensor integration is paramount for the system performance, *i.e.*, to accurately estimate both position and orientation. This makes it mandatory to use a sensor fusion algorithm (such as a Kalman filter or

a similar one [12,18–20]) and to be also very careful with data formats and with the definition of the reference frames (especially with GPS measurements from different devices) [21].

To start with, we have to define the different local reference frames, as well as their relations in terms of translations and rotations [20]. In this way, the origins of four frames will be defined, O_{TOP} , O_{IMU} , O_{CAM} and O_{PRI} , corresponding to the Topcom GNSS, IMU, camera and prism, respectively. However, to define the frame, orientation is also needed. By default, the IMU local reference frame is aligned to its external housing, with X_{IMU} pointing forward, Y_{IMU} pointing left and Z_{IMU} pointing upward, as shown in Figure 4. Taking into account that orientation with respect to the global world frame is provided by the IMU, sensors were placed within the box in such a way that rotations were avoided. Therefore, the orientation of frames TOP , PRI and CAM was defined as having the same orientation as the IMU frame. Finally, translation vectors from a reference frame to another have to be defined and accurately measured. For instance, Figure 4 depicts vector d_{IMU}^{TOP} , which stands for the position of IMU in the reference frame, TOP .

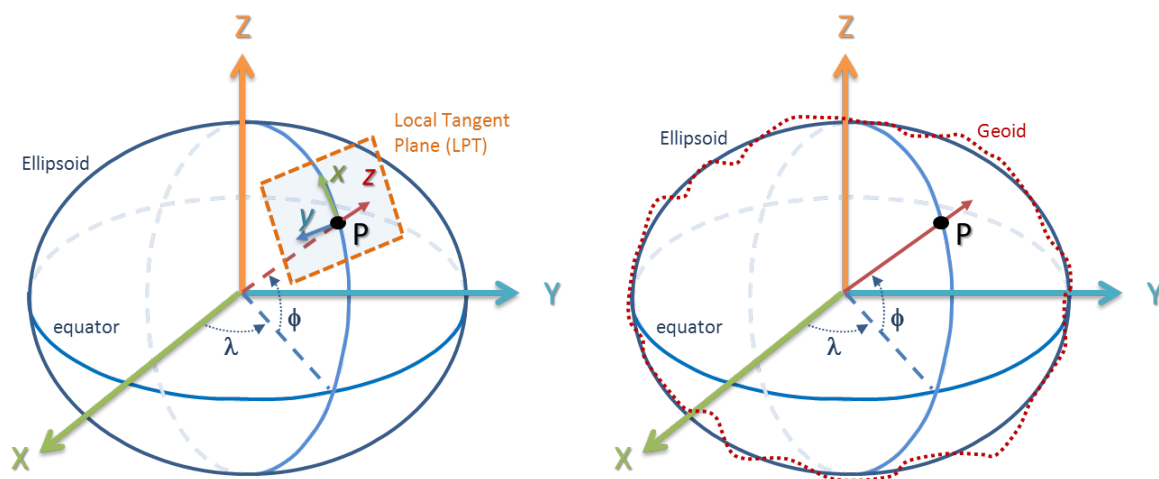
Figure 4. Local reference frames.



The next step is to select a global reference frame. Since we are dealing with positioning, it is preferred to use an Earth-Centered Earth-Fixed (ECEF) coordinate system that rotates with the Earth [21]. In the ECEF system, the XY plane coincides with the equator, the Z -axis points to the north and the X -axis points to the Greenwich meridian (0° longitude), as shown in Figure 5. However, it is difficult to relate global coordinates of a point in an XYZ ECEF reference system with a map of Earth's surface. To solve this problem, it is possible to use a physical model of the Earth, such as the World Geodetic System 1984 (WGS84), which is the standard for GPS applications, to define a point in terms of latitude, longitude and altitude, *i.e.*, in a spherical coordinate system [21].

In WGS84, Earth's shape is approximated by an ellipsoidal model defined by several parameters (semi-major and semi-minor axes, inverse flattening, *etc.*). Because this could be rather generic, accuracy can be improved by using local models (called datums), which are, basically, different values for the ellipsoid axes and center [20]. The GPS receiver integrated with the IMU uses the generic WGS84 datum.

Figure 5. The Earth-Centered Earth-Fixed (ECEF) coordinate system and Local Tangent Plane (left) and ellipsoid vs. Earth geoid (right).



Even more, the spherical coordinate system of WGS84 can be transformed into a Cartesian coordinate system by means of a linearization of the WGS84 ellipsoid, yielding a local tangent plane (LTP) (see Figure 5), with X_{LTP} pointing north, Y_{LTP} pointing west and Z_{LTP} pointing upward.

Another, commonly used Cartesian coordinate system is the Universal Transverse Mercator (UTM), which is also based on the WGS84 ellipsoid to model the Earth [21]. UTM divides the surface of the Earth into sixty zones (called UTM zones) and gives locations in a 2D Cartesian coordinate system, with the X -axis pointing east and the Y -axis pointing north. It is also known as the easting-northing coordinate system. Notice that UTM implies a 90-degree rotation around Z with respect to LTP, which is a northing-westing description.

Finally, it should be remarked that GPS altitude can be measured above the ellipsoid or above mean sea level (MSL) [20]. Both altitudes can be quite different, since there are significant differences between the ellipsoid and Earth geoid (see Figure 5 (right)). By default, Xsens GPS provides altitude above the ellipsoid. However, Topcon GR-3 GNSS uses the National Marine Electronics Association (NMEA) standard, which deals with altitude above the geoid. Fortunately, it also provides the geoidal separation, *i.e.*, the difference between the Earth ellipsoid and geoid defined by the reference datum. Altitude above the ellipsoid can be then obtained by adding MSL altitude and geoidal separation (which can be negative). In this system, NMEA altitude will be used as reference, so geoidal separation will be subtracted from the Xsens GPS altitude. In the experiments shown in Section 3, the geoidal separation was 48.776 m.

On the other hand, we use the IMU to measure a global orientation. The problem here is that inertial measurements should be referred to as an Earth Centered Inertial (ECI) coordinate system, which does

not rotate with the Earth, instead of referred to as an ECEF, as the GPS position does. Fortunately, it is possible to work in the local linearized tangent plane (LTP) without significant errors. Therefore, the LTP can be used as the global coordinate system to define both position and orientation, with X_{LTP} pointing north, Y_{LTP} pointing west and Z_{LTP} pointing upward.

Therefore, IMU provides the rotation of IMU reference system with respect to the LTP . This rotation is given in terms of Euler angle roll (ϕ), pitch (θ) and yaw/heading (ψ), to perform subsequent rotations in the $X_{LTP}Y_{LTP}Z_{LTP}$ axes, which can be expressed by means of the Direction Cosine Matrix (DCM):

$$\begin{aligned} R_{IMU}^{LTP} &= R_{\psi}^{Z_{LTP}} R_{\theta}^{Y_{LTP}} R_{\phi}^{X_{LTP}} \\ &= \begin{pmatrix} \cos \theta \cos \psi & \sin \phi \sin \theta \cos \psi - \cos \phi \sin \psi & \cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi \\ \cos \theta \sin \psi & \sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi & \cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi \\ -\sin \theta & \sin \phi \cos \theta & \cos \phi \cos \theta \end{pmatrix} \quad (1) \end{aligned}$$

Notice that the DCM of Equation (1) will also provide the global rotation of the TOP , PRI and CAM reference systems, because they were defined to have the same orientation as IMU .

It is well known that Euler angles present a singularity when the pitch angle (θ) is close to $\pm 90^\circ$ [20]. In our system, this would imply that X_{IMU} is pointing upward or downward, *i.e.*, the box would be in the horizontal position (see Figure 4), which is not possible. Therefore, there is no need to use an alternative representation, such as quaternion.

In this way, getting the global position of TOP from the GNSS and the Euler angles from the IMU, it is possible to compute, for instance, the global position of the camera, O_{CAM} , by means of the next transformation (camera orientation is given by the Euler angles):

$$O_{CAM}^{LTP} = O_{TOP}^{LTP} + R_{TOP}^{LTP} \cdot d_{CAM}^{TOP} \quad (2)$$

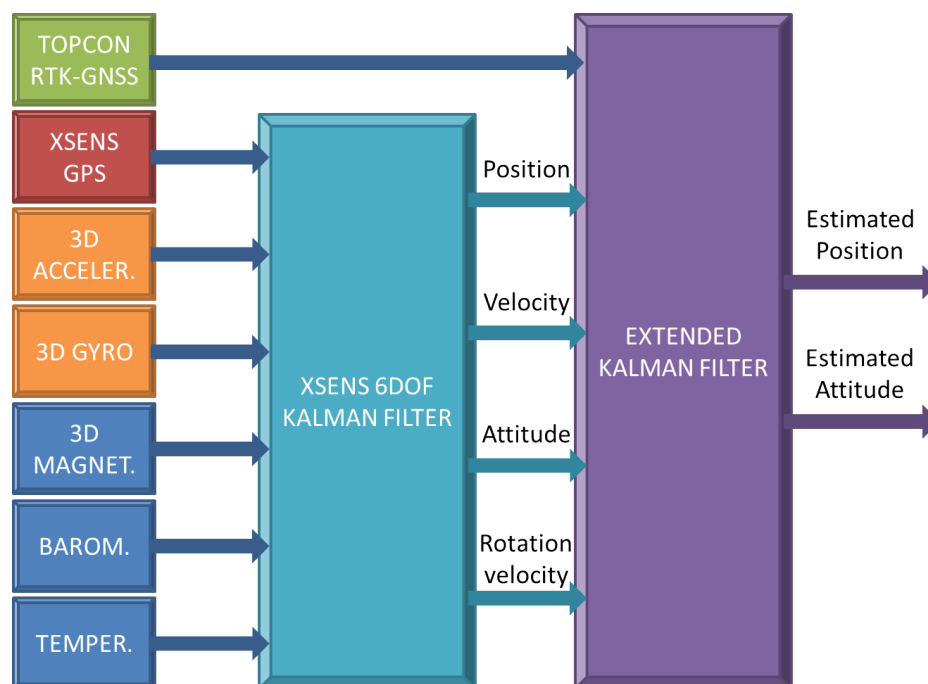
Once the sensors measurements have been adapted and transformed into a common global reference system, the next step is to perform sensor fusion to improve accuracy and reliability. To that end, a Kalman filter approach can be used [12,18–20,22]. Basically, a Kalman filter corrects measurements by comparing them with their predicted values, but at the same time, predictions are corrected based on the actual measurements [20]. The idea behind this is that both estimations and observations have noise or errors. Then, a prediction model, based on the system equations, as well as an observation model have to be defined. Firstly, in the prediction step, state system variables are integrated based on the prediction model to obtain the state estimates. Secondly, in the correction step, state estimation is updated, taking into account the actual observations. These steps are repeated in a cycle.

Given the nonlinear nature of the proposed system, an extended Kalman filter (EKF) will be preferred to a linear Kalman filter. Furthermore, a cascaded EKF has been implemented (see Figure 6) [20]. The reasons for this configuration are threefold: first, to take advantage of the built-in Xsens Kalman filter to process inertial measurements; second, the design is more simple, and it is easier to tune; and third, even though the cascade filter is suboptimal, its performance is comparable to a full EKF [19].

The first Kalman filter is mainly devoted to integrating inertial sensors (orange blocks in Figure 6) to obtain the position and orientation estimates (prediction step). Without corrections, small errors

and inaccuracies will cause the estimation errors to grow with time. Thus, in the correction step, the filter uses the stable position and velocity from the GPS (red block) to correct the estimation drift. Moreover, the orientation can also be improved from the GPS, since it is used to derive the position (see Equation (2)). Therefore, it is a loosely-coupled Kalman filter, since GPS data is used to correct IMU estimation. Finally, blue blocks are used as additional aiding sensors in the correction step. For instance, 3D magnetometers can be used to improve true north estimation based on Earth's magnetic field. The barometer is also used as an aiding sensor to measure altitude changes, increasing the vertical accuracy.

Figure 6. The Kalman filter structure to obtain system pose estimation (position and orientation).



It should be noted that, as stated above, the built-in GPS has poor accuracy compared with the Topcon GR-3 RTK-GNSS, so it could seem convenient to provide corrections to the Xsens GPS to be propagated into the filter. However, it is not allowed by the device. Therefore, the second filter is mainly designed to take advantage of the high accuracy of the Topcon GR3 RTK-GNSS to estimate the system pose (position and orientation).

The EKF is based on the one proposed by Gavrillets (see [23] for details). EKF requires one to compute the derivative of the rotation matrix, which cannot be solved in closed form when using Euler angles [20]. Therefore, using the quaternion method within the EKF will be more effective. The quaternion is free of singularities, but is subject to a unit norm constraint [23]. Then, the state vector is defined as:

$$x = [p \ v \ q \ a_b \ \omega_b] \quad (3)$$

where $p = [x \ y \ z]^T$ is the position, $v = [v_x \ v_y \ v_z]^T$ the velocity, $q = [q_0 \ q_1 \ q_2 \ q_3]^T$ is a unit norm quaternion representation of orientation, $a_b = [a_{x_b} \ a_{y_b} \ a_{z_b}]^T$ is the vector of acceleration bias and $\omega_b = [\dot{\phi}_b \ \dot{\theta}_b \ \dot{\psi}_b]^T$ is the vector of gyro rate bias.

It should be noted that quaternion and Euler angles are related by means of the next equations:

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \cos(\phi/2) \cos(\theta/2) \cos(\psi/2) + \sin(\phi/2) \sin(\theta/2) \sin(\psi/2) \\ \sin(\phi/2) \cos(\theta/2) \cos(\psi/2) - \cos(\phi/2) \sin(\theta/2) \sin(\psi/2) \\ \cos(\phi/2) \sin(\theta/2) \cos(\psi/2) + \sin(\phi/2) \cos(\theta/2) \sin(\psi/2) \\ \cos(\phi/2) \cos(\theta/2) \sin(\psi/2) - \sin(\phi/2) \sin(\theta/2) \cos(\psi/2) \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} \phi \\ \theta \\ \psi \end{bmatrix} = \begin{bmatrix} \tan^{-1} \left(\frac{2q_2q_3 + 2q_0q_1}{2q_0^2 + 2q_3^2 - 1} \right) \\ \sin^{-1} (2q_0q_2 - 2q_1q_3) \\ \tan^{-1} \left(\frac{2q_1q_2 + 2q_0q_3}{2q_0^2 + 2q_1^2 - 1} \right) \end{bmatrix} \quad (5)$$

The discrete-time update equations are:

$$\begin{aligned} p_{t+1} &= p_t + \dot{p}_t \cdot dt \\ v_{t+1} &= v_t + \dot{v}_t \cdot dt \\ q_{t+1} &= \exp \left(-\frac{1}{2} \Omega \cdot dt \right) q_t \\ a_{bt+1} &= a_{bt} + \eta_{a_{bt}} \\ \omega_{bt+1} &= \omega_{bt} + \eta_{\omega_{bt}} \end{aligned} \quad (6)$$

where η_{a_b} and η_{ω_b} are noise terms and Ω is a skew-symmetric matrix composed of the gyro rate measurements and bias:

$$\Omega = \begin{bmatrix} 0 & (\dot{\phi} - \dot{\phi}_b) & (\dot{\theta} - \dot{\theta}_b) & (\dot{\psi} - \dot{\psi}_b) \\ -(\dot{\phi} - \dot{\phi}_b) & 0 & -(\dot{\psi} - \dot{\psi}_b) & (\dot{\theta} - \dot{\theta}_b) \\ -(\dot{\theta} - \dot{\theta}_b) & (\dot{\psi} - \dot{\psi}_b) & 0 & -(\dot{\phi} - \dot{\phi}_b) \\ -(\dot{\psi} - \dot{\psi}_b) & -(\dot{\theta} - \dot{\theta}_b) & (\dot{\phi} - \dot{\phi}_b) & 0 \end{bmatrix} \quad (7)$$

Corrections for the Topcon GR3 RTK-GNSS, for instance, can be done based on the next lever arm equations for position and velocity [22]:

$$O_{TOP}^{LTP} = p + R \cdot d_{IMU}^{TOP} + \eta_p \quad (8)$$

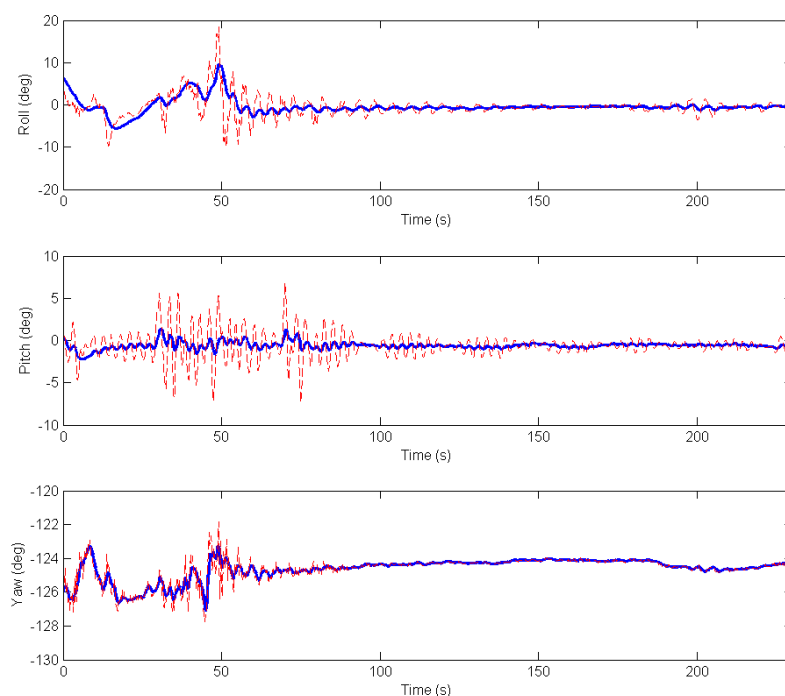
$$V_{TOP}^{LTP} = v + R \cdot \omega \times d_{IMU}^{TOP} + \eta_v$$

where η_p and η_v are noise terms, ω is the gyro rate vector, $[\dot{\phi}, \dot{\theta}, \dot{\psi}]$, and R is the DCM of the orientation angles that in terms of the quaternion is given by:

$$R = \begin{bmatrix} 2q_0^2 + 2q_1^2 - 1 & 2q_1q_2 - 2q_0q_3 & 2q_1q_3 + 2q_0q_2 \\ 2q_1q_2 + 2q_0q_3 & 2q_0^2 + 2q_2^2 - 1 & 2q_2q_3 - 2q_0q_1 \\ 2q_1q_3 - 2q_0q_2 & 2q_2q_3 + 2q_0q_1 & 2q_0^2 + 2q_3^2 - 1 \end{bmatrix} \quad (9)$$

The accuracy of the estimation will be validated in the next section using a robotic total station in field experiments. Figure 7 shows the output of the filter for the orientation angles during four minutes after deployment of the system.

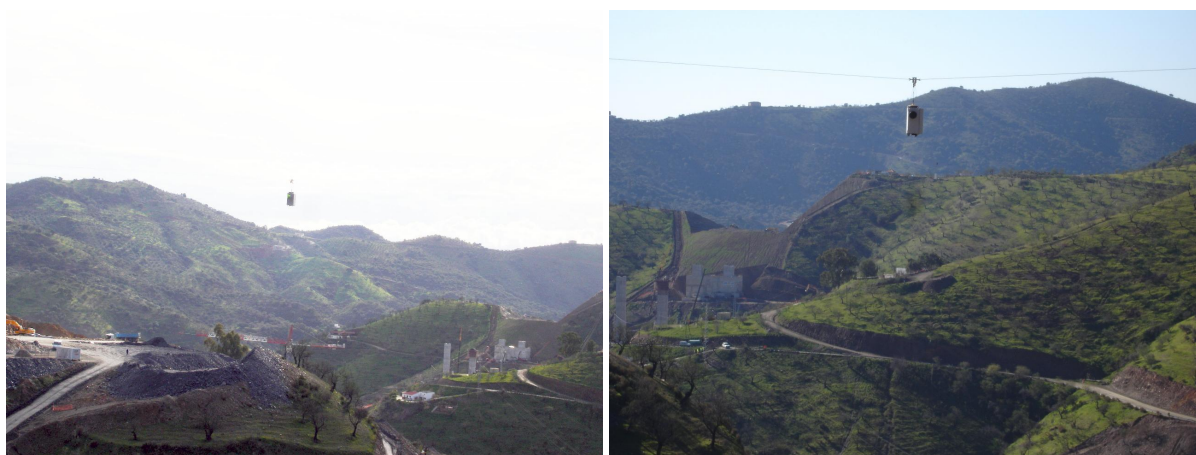
Figure 7. Roll, pitch and yaw filtered (blue) by the extended Kalman filter (EKF) during four minutes after deployment.



3. Experimental Results

Several experiments have been conducted to test the system. These experiments were performed in a highway construction site (see Figure 8) and included, among others, the validation of the estimated position and orientation using a robotic total station, the computation of a digital surface model (DSM) and georeferencing based on camera position estimation or ground control points.

Figure 8. Field experiments site.



Thus, in the first experiment, seven points, located along the system trajectory, were measured with a robotic total station pointing to the on-board prism, as shown in Figure 9. The coordinates from the total station, defined in a local reference frame, were translated into UTM and compared with the prism

coordinates estimated by the cascaded EKF, as shown in Table 1. The Root Mean Squared Error (RMSE) was 0.04 m and the Mean Absolute Error (MEA), 0.03 m. Notice that orientation estimations are also used to compute the prism coordinates, due to the lever arm effect.

Figure 9. Robotic total station and measured points (red) vs. estimated (green) along the trajectory.

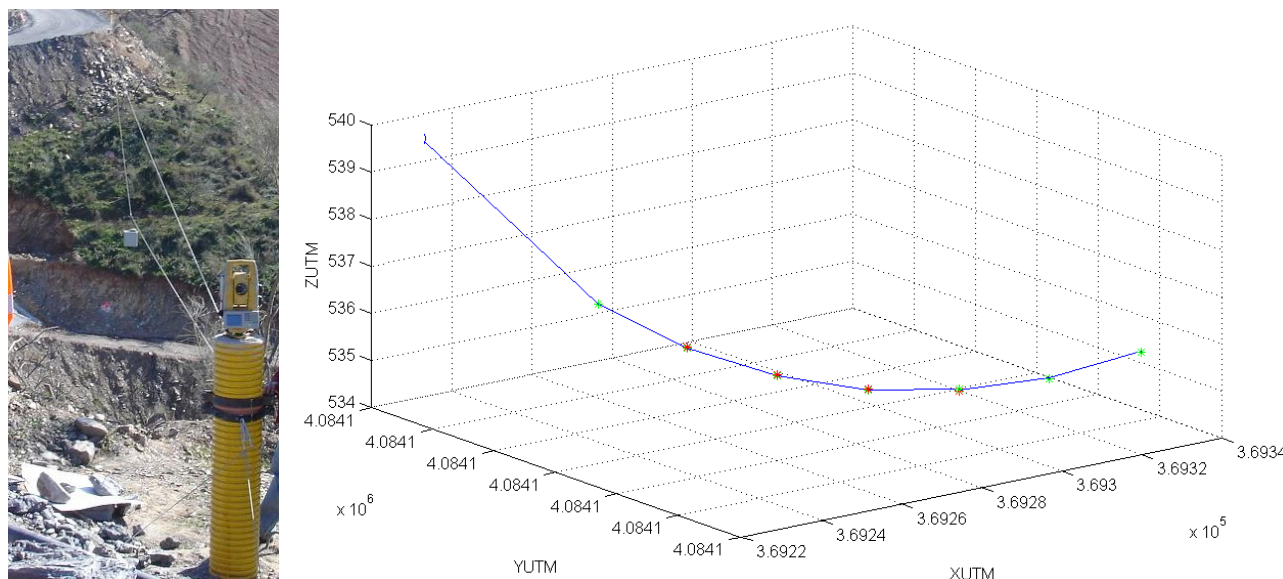


Table 1. Position of the on-board prism measured with the total station (TS) in Universal Transverse Mercator (UTM) coordinates, estimated by the Kalman filter (EKF) and error.

Point	X_{TS}	Y_{TS}	Z_{TS}	X_{EKF}	Y_{EKF}	Z_{EKF}	Error (m)
1	369,251.77	4,084,123.89	536.37	369,251.77	4,084,123.89	536.37	0.01
2	369,263.71	4,084,117.20	535.56	369,263.72	4,084,117.20	535.54	0.03
3	369,275.59	4,084,110.31	535.08	369,275.60	4,084,110.32	535.05	0.03
4	369,287.86	4,084,103.47	534.88	369,287.84	4,084,103.43	534.86	0.05
5	369,299.87	4,084,096.54	534.93	369,299.88	4,084,096.57	534.97	0.06
6	369,311.94	4,084,089.74	535.29	369,311.94	4,084,089.73	535.30	0.02
7	369,324.05	4,084,082.69	535.97	369,324.06	4,084,082.71	535.98	0.03

A second experiment was devoted to collecting images in order to construct a digital surface model (DSM), *i.e.*, a model of the Earth's surface, including all objects on it, of a rough, bumpy and sloping terrain. This type of model is, sometimes, also known as a digital elevation model (DEM), while digital terrain models (DTMs) focus on the bare ground surface without any objects. There are some, very accurate, commercial applications, like the Leica Photogrammetry Suite (LPS) [24], to compute such models. However, in this paper, instead of commercial software, we made use of open applications to compute the DSM based on the emergent structure from motion (SfM) techniques [25–33].

Structure from motion intends to estimate a 3D structure based on a sequence of 2D images (see Figure 10). Moreover, with these techniques, it is possible to estimate both the inner and exterior

orientation of the camera. The inner orientation corresponds to the camera geometry (it is defined by parameters, such as the focal length, principle point, lens distortions, *etc.*), while the exterior orientation is its position and orientation when the image was captured. Another interesting feature of these techniques is that they can be easily applied by means of open software, as will be shown below.

Figure 10. Image set used in the structure from motion approach.



Some of these open software applications are mainly devoted to the determination of the camera parameters and a sparse point cloud reconstruction (VisualSFM [25], Bundler [28], APERO [31], *etc.*) while others provide dense point clouds or polygonal models (Patch-based Multi-view Stereo Software (PMVS2) [27], MicMac [32], SURE [33]...). In [34], Remondino *et al.* present a very interesting, in depth comparison of these tools based on different large and complex datasets. In this paper, we have used VisualSFM and PMVS2. VisualSFM offers a GUI for image feature extraction, automatic matching, ground control points management and interfacing with PMVS2.

Figure 10 depicts the 25 image sequence that will be used to compute the DSM. In the first step, the software automatically extracts features that can be of interest in each image, as shown in Figure 11. Feature detection is based on the Scale Invariant Feature Transform (SIFT) [30], which is a region detector, robust with respect to rotations, translations or scales in images [35].

Secondly, an automatic matching of features in every pair of images is performed. Figure 12 shows the matching of a single feature (left) and all the features matched in a given pair of images (right). From the matched points in different images, a Bundle Block Adjustment [26,36] is performed in order

to compute the inner and external orientations of the camera, which minimize the re-projection error. Generally, it is formulated as a nonlinear least-squares problem, where the error is the squared norm of the difference between the observed feature location and the projection of the corresponding 3D point on the image plane of the camera and numerically solved by using the Levenberg–Marquardt (LM) algorithm [37]. Moreover, VisualSFM implements a Multicore Bundle Adjustment [36] that exploits hardware parallelism for efficiently solving large-scale 3D scene reconstruction problems.

Figure 11. Features of interest detected in an image.

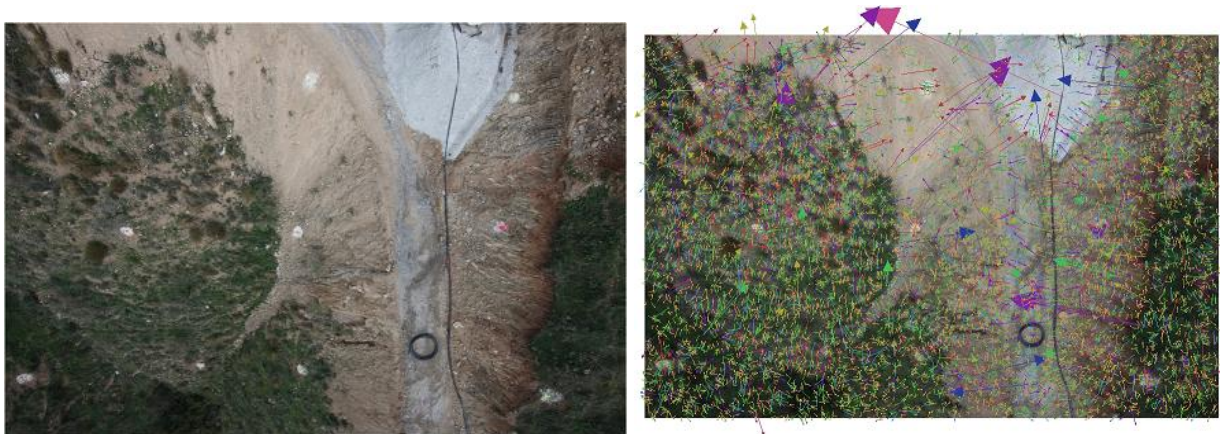
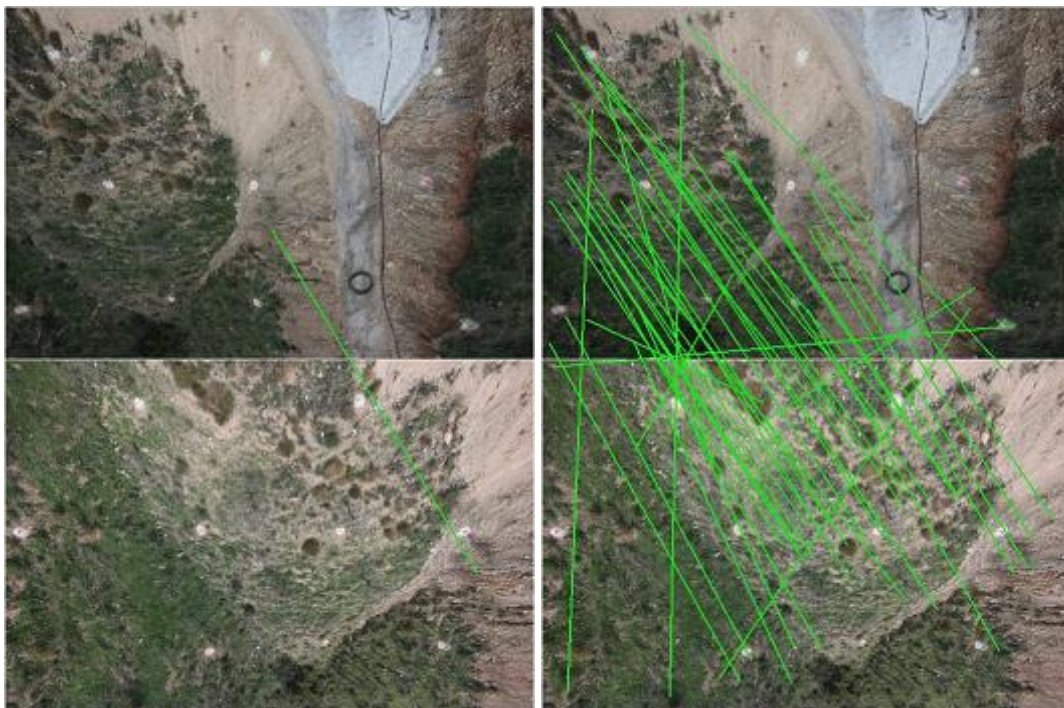


Figure 12. Single feature matching (**left**) and the complete set of matchings (**right**) in a given pair of images.



In this way, it is possible to obtain a 3D representation of the matched points, which is a sparse point cloud, as shown in Figure 13. This figure also depicts the 25 camera pose estimates. The sparse point cloud has 26.659 points (13.556 of them seen by more than three cameras) and 90.609 projections

(65.501 seen by more than three cameras). The adjustment estimated three translations, three rotations, the focal length and one radial distortion parameter for each image, with a residual error of 1.41.

Figure 13. Sparse point cloud and estimated camera locations.



Moreover, it is also possible to compute a dense point cloud from the sparse one. Such a computation is performed by the Patch-based Multi-view Stereo Software (PMVS2) based on a multi-view stereo algorithm [27]. Figure 14 shows the obtained dense point cloud with a 1.217.381 vertex (the sparse point cloud had 26.659 points). The new dense point cloud can be converted into a surface by means of, for instance, a Poisson reconstruction algorithm [38], yielding a digital surface model (DSM) of the rough, bumpy and sloping terrain, as the one shown in Figures 15 and 16.

Figure 14. Dense point cloud representation of the experiment site.



Nevertheless, the computed DSM is defined within its local coordinate system. Therefore, the final step consists of georeferencing the model, *i.e.*, the model coordinates have to be transformed into world coordinates. This can be performed, for instance, by applying Horn's method [39] to compute the optimal affine transformation (rotation R , translation T and scaling S) that best maps a set of points of the model, P^M , into the real world, P^W :

$$P^W = S \cdot R \cdot P^M + T \quad (10)$$

Figure 15. Different views of the computed digital surface model.

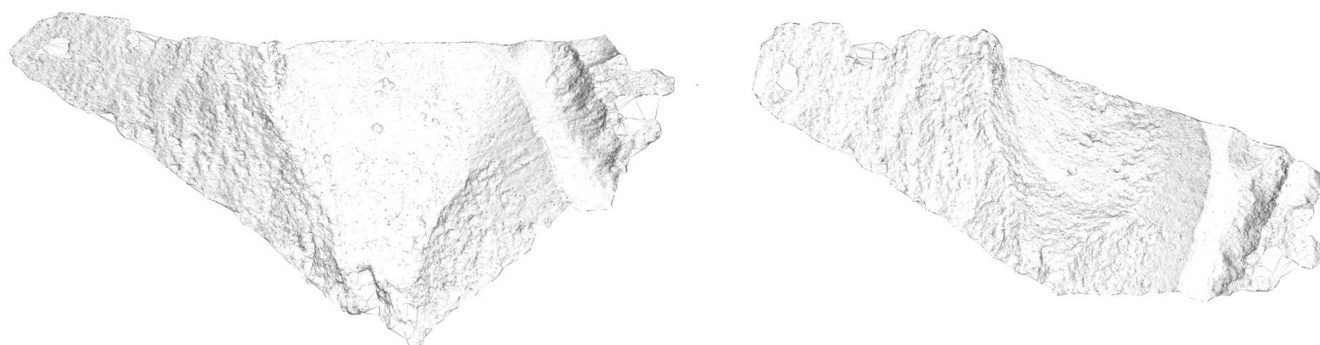


Figure 16. Different views of the textured digital surface model.



Obviously, both model and world coordinates need to be known for each point in the set. To that end, it is possible to use camera coordinates or ground control points. Georeferencing by using camera coordinates measured by the system is called direct georeferencing. In this case, coordinates (in the local model and the real world) of 25 cameras were available to compute the transformation of Equation (10). Table 2 shows the camera coordinates in the local model, the ones derived from the EKF (in UTM) and the world coordinates resulting from applying the affine transformation, as well as the norm of the difference between both positions. The Root Mean Squared Error (RMSE) considering the 25 cameras is 0.23 m and the Mean Absolute Error (MAE), 0.20 m.

Table 2. Camera coordinates in the local model (L), UTM estimated by the Kalman filter (EKF), results from the transformation (T) and error.

X_L	Y_L	Z_L	X_{EKF}	Y_{EKF}	Z_{EKF}	X_T	Y_T	Z_T	Error (m)
7.954	0.007	−0.162	369,319.94	4,084,085.17	535.76	369,319.89	4,084,085.23	535.52	0.25
6.381	0.029	−0.122	369,315.74	4,084,087.54	535.52	369,315.68	4,084,087.55	535.33	0.20
4.884	0.024	−0.109	369,311.74	4,084,089.84	535.34	369,311.71	4,084,089.83	535.23	0.11
3.537	0.007	−0.112	369,308.14	4,084,091.87	535.21	369,308.16	4,084,091.91	535.19	0.05
2.255	−0.043	−0.115	369,304.68	4,084,093.89	535.13	369,304.83	4,084,093.98	535.14	0.18
1.045	−0.047	−0.132	369,301.51	4,084,095.93	535.04	369,301.62	4,084,095.82	535.14	0.19
0.019	−0.027	−0.129	369,298.99	4,084,097.41	535.00	369,298.86	4,084,097.32	535.09	0.18
...

Figure 17. Ground control points in the experiment area.

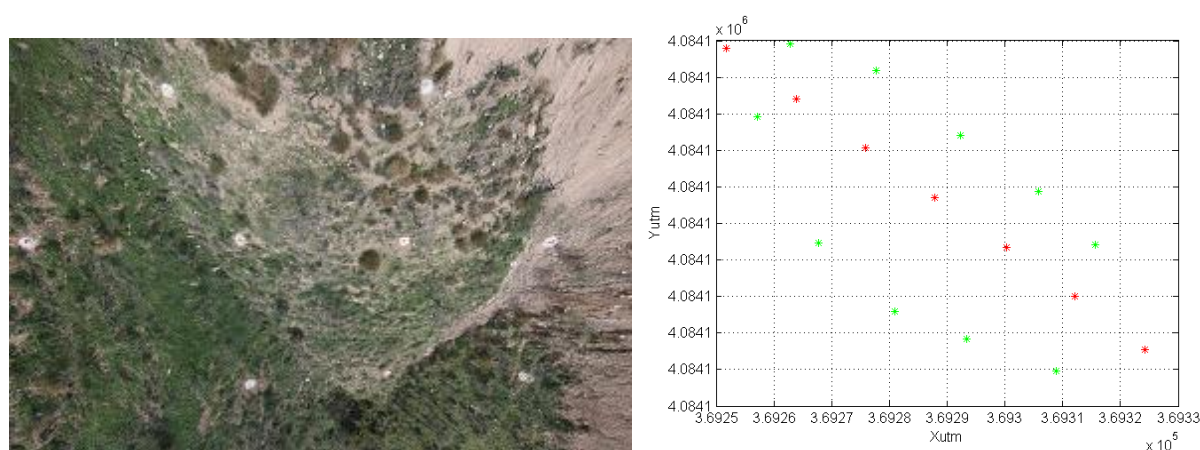
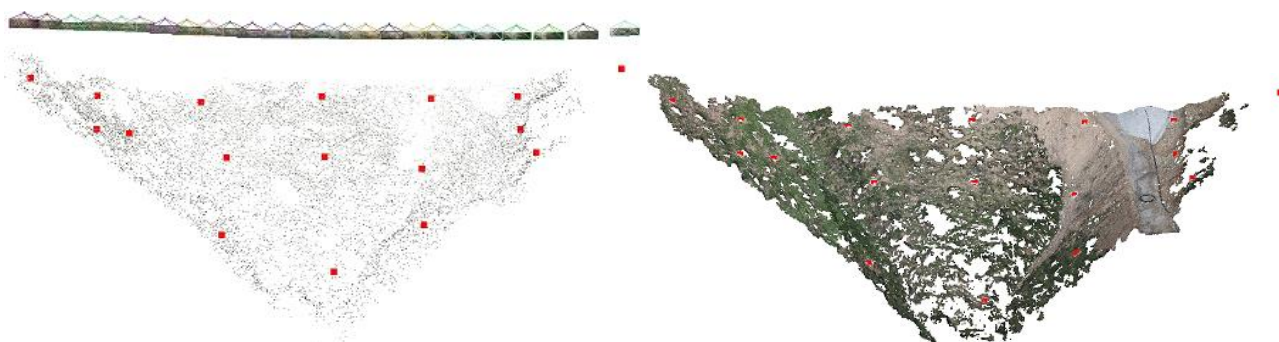


Figure 18. Ground control points on the sparse and dense point clouds.



Finally, another option is georeferencing based on ground control points, *i.e.*, by transforming the GCP coordinates in the local model into the GCPs in the real world. To improve performance, GCPs have to be distributed in such a way that several of them appear in consecutive overlapping images. Seventeen GCPs were signaled in the experiment area (see Figure 17), by red (7) or yellow (10) circular targets on a white background, and their coordinates recorded with a total station. VisualSFM facilitates GCP

management. Thus, GCP can be manually matched directly onto the 2D images and, then, automatically converted to the 3D model, as shown in Figure 18. Then, the affine transformation that converts the seventeen GCPs from model coordinates into world coordinates is calculated, yielding approximation errors of 0.17 m RMSE and 0.13 m MAE.

4. Conclusions

Advances in sensor technologies have provided surveyors and construction engineers with a complete set of high accuracy ground-level tools (robotic total stations, TLS, RTK-GNSS, *etc.*) for surveying, layout and grade management of construction projects. On the other hand, improvements in digital image, CMOS camera sensors, MEMS inertial measurement units (IMU), integrated IMU-GPS devices, laser sensors, *etc.*, have produced lightweight, low-cost MEMS-based sensors that have boosted the spread of unmanned aircraft systems applied to aerial photogrammetry, surveying and engineering applications.

The counter-part of these low-cost devices is a reduction of accuracy and performance, which can be mitigated by using sensor fusion techniques for pose estimation. Moreover, these vehicles usually have a reduced payload (lower than 3 kg), limited autonomy and even require skilled operators, either to remote control or as safety pilots in the case of UAVs with autopilot systems (in some cases, the autopilot can only be activated when the vehicle is in the air, but take-off and landing have to be operated manually). UAVs without these limitations have a significantly higher cost.

This paper has presented a new system designed in a research project, funded by the Spanish construction company, Sacyr SAU, to take advantage of ground device accuracy (facilitating integration with surveying total stations, RTK-GNSS and the construction site infrastructure) and, at the same time, overcoming the limitations of aerial vehicles with respect to weight, autonomy and the need for skilled operators. Thus, the system integrates an accurate RTK-GNSS, commonly used at construction sites, a prism for interfacing with a robotic total station, a high resolution camera and a medium-cost IMU-GPS, into a device, weighing around 12 kg in total. It moves between two mounting points, in a blondin ropeway configuration, at the construction site, taking pictures and recording position and orientation data along the cable path. Using two mounting points can be a disadvantage with respect to the coverage area, but this could be solved by using a Skycam configuration.

A cascaded extended Kalman filter is used to integrate measurements from the on-board inertial measurement unit (IMU), a GPS and a RTK-GNSS, which has been validated with a robotic total station in field experiments in a construction site, with an error of 0.03 m (MAE).

Finally, instead of commercial software, in this paper, we have made use of open software to compute a digital surface model (DSM) of a rough, bumpy and sloping terrain. Namely, structure from motion (SfM) techniques have been applied to obtain the 3D model from a sequence of 25 images. Among the available software, VisualSfM was selected, because its GUI facilitates image feature extraction, automatic matching, ground control points management and interfacing with PMVS2 (for dense point cloud calculation). The Bundle Block Adjustment resulted in a sparse point cloud with 26.659 points (13.556 of them seen by more than three cameras). On the other hand, the obtained dense point cloud had a 1.217.381 vertex. The georeferencing of the DSM was performed based on the estimated camera

positions, with an approximation error of 0.20 m (MAE), and also using 17 ground control points (GCP), with an error of 0.13 m (MAE). Even if these results are not accurate enough for certain applications, they can be considered good, because the slopes of the terrain and the fact that the software was used as is, without particular improvements. However, this should be tested with larger sequences of images and different terrains.

Acknowledgments

This work has been supported by the Spanish construction company, Sacyr SAU. Acknowledgments are due to F. López Pichaco, C. Nogales and A. Gallego for their collaboration in the construction of the prototype, as well as in the realization of field experiments. The authors would also like to thank the reviewers for their time and their valuable comments.

References

1. Lloret, P. Inertial+total Station+GPS for High Productivity Surveying , In Proceedings of the Position Location and Navigation Symposium, Las Vegas, NV, USA, 20–23 March 1990; pp. 338–346.
2. Gikas, V. Three-dimensional laser scanning for geometry documentation and construction management of highway tunnels during excavation. *Sensors* **2012**, *12*, 11249–11270.
3. Park, H.S.; Son, S.; Choi, S.W.; Kim, Y. Wireless laser range finder system for vertical displacement monitoring of mega-trusses during construction. *Sensors* **2013**, *13*, 5796–5813.
4. Siu, M.F.; Lu, M.; AbouRizk S. Combining photogrammetry and robotic total stations to obtain dimensional measurements of temporary facilities in construction field. *Vis. Eng.* **2013**, doi:10.1186/2213-7459-1-4.
5. El-Omari, S.; Moselhi, O. Integrating 3D laser scanning and photogrammetry for progress measurement of construction work. *Autom. Constr.* **2008**, *18*, 1–9.
6. Fregonese, L.; Barbieri, G.; Biolzi, L.; Bocciarelli, M.; Frigeri, A.; Taffurelli, L. Surveying and monitoring for vulnerability assessment of an ancient building. *Sensors* **2013**, *13*, 9747–9773.
7. Eisenbeiss, H. UAV Photogrammetry. Doctoral Thesis, Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland, 2009; p. 235.
8. Skaloud, J.; Vallet, J.; Keller, K.; Veyssi re, G.; K l bl, O. An eye for landscapes: Rapid aerial mapping with handheld sensors. *GPS World* **2006**, *17*, 26–32.
9. Peterman, V.; Mesari , M. Land Survey from Unmanned Aerial Vehicle. In Proceedings of International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B1, 2012 XXII ISPRS Congress, Melbourne, Australia, 25 August–1 September 2012; pp. 447–452.
10. Dorninger, P.; Pfeifer, N. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. *Sensors* **2008**, *8*, 7323–7343.

11. Manyoky, M.; Theiler, P.; Steudler, D.; Eisenbeiss, H. Unmanned Aerial Vehicle in Cadastral Applications. In Proceedings of International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVIII-1/C22 UAV-g 2011, Conference on Unmanned Aerial Vehicle in Geomatics, Zurich, Switzerland, 14–16 September 2011; pp. 1–6.
12. Wallace, L.; Lucieer, A.; Watson, C.; Turner, D. Development of a UAV-LiDAR system with application to forest inventory. *Remote Sens.* **2012**, *4*, 1519–1543.
13. Chiabrando, F.; Nex, F.; Piatti, D.; Rinaudo, F. UAV and RPV systems for photogrammetric surveys in archaeological areas: Two tests in the Piedmont region (Italy). *J. Archaeol. Sci.* **2011**, *38*, 697–710.
14. Turner, D.; Lucieer, A.; Watson, C. An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (UAV) imagery, based on structure from motion (SfM) point clouds. *Remote Sens.* **2012**, *4*, 1392–1410.
15. Aber, J.S.; Aber, S.W.; Pavri, F. Unmanned Small-format Aerial Photography from Kites for Acquiring Large-scale, High-resolution, Multiview-angle Imagery. In Proceedings of the Pecora. 15/Land Satellite Information IV/ISPRS Commission (I/FIEOS 2002), Denver, CO, USA, 10–14 November 2002.
16. Watts, A.C.; Ambrosia, V.G.; Hinkley, E.A. Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. *Remote Sens.* **2012**, *4*, 1671–1692.
17. Nex, F.; Remondino, F. UAV for 3D mapping applications: A review. *Appl. Geomat.* **2013**, 1–15.
18. Chao, H.; Coopmans, C.; Di, L.; Chen, Y.Q. A Comparative Evaluation of Low-cost IMUs for Unmanned Autonomous Systems. In Proceedings of IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems, Salt Lake City, UT, USA, 5–7 September 2010; pp. 211–216.
19. Kingston, D.B.; Beard, R.W. Real-time Attitude and Position Estimation for Small UAVs Using Low-cost Sensors. In Proceedings of AIAA 3rd Unmanned Unlimited Technical Conference, Workshop and Exhibit, Chicago, IL, USA, 21–23 September 2004; pp. 2004–6488.
20. Noureldin, A.; Karamat, T.B.; Georgy, J. *Fundamentals of Inertial Navigation, Satellite-based Positioning and Their Integration*; Springer: Berlin/Heidelberg, Germany, 2013.
21. Hofmann-Wellenhof, B.; Lichtenegger, H.; Wasle, E. *GNSS Global Navigation Satellite Systems*; Springer: Berlin/Heidelberg, Germany, 2008.
22. Seo, J.; Lee, H.K.; Lee, J.G.; Park, C.G. Lever arm compensation for GPS/INS/Odometer integrated system. *Int. J. Control Autom. Syst.* **2006**, *4*:2, 247–254.
23. Gavrillets, V. Autonomous Aerobatic Maneuvering of Miniature Helicopters. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2003.
24. Wang, Y.; Yang, X.; Stojic, M.; Skelton, B. Toward Higher Automation and Flexibility in Commercial Digital Photogrammetric System. In Proceedings of International Archives of Photogrammetry and Remote Sensing, Volume IC WG II/IV, Istanbul, Turkey, 12–23 July 2004; pp. 838–840.
25. Wu, C. VisualSfM: A Visual Structure from Motion System, 2011. Available online: <http://ccwu.me/vsfm/> (accessed on 11 October 2013).
26. Snavely, N.; Seitz, S.M.; Szeliski, R. Modeling the world from internet photo collections. *Int. J. Comput. Vis.* **2008**, *80*, 189–210.

27. Furukawa, Y.; Ponce, J. Accurate, dense, and robust multi-view stereopsis. *IEEE Trans. Pattern Anal.* **2009**, *32*, 1362–1376.
28. Snavely, N. Bundler: Structure from Motion (SfM) for Unordered Image Collections. Available online: phototour.cs.washington.edu/bundler/ (accessed on 11 October 2013).
29. Wu, C. Towards Linear-time Incremental Structure from Motion. In Proceedings of 2013 International Conference on 3DTV-Conference, Seattle, WA, USA, 29 June–1 July 2013; pp. 127–134.
30. Wu, C. SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT). 2007. Available online: cs.unc.edu/~ccwu/siftgpu/ (accessed on 11 October 2013).
31. Pierrot-Deseilligny, M.; Cléry, I. APERO, an Open Source Bundle Adjustment Software for Automatic Calibration and Orientation of a Set of Images. In Proceedings of International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-5/W16, Trento, Italy, 2–4 March 2011.
32. Pierrot-Deseilligny, M.; Paparoditis, N. A Multiresolution and Optimization-based Image Matching Approach: An Application to Surface Reconstruction from SPOT5-HRS Stereo Imagery. In Proceedings of International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI-1/W41, Ankara, Turkey, 14–16 February 2006.
33. Rothermel, M.; Wenzel, K.; Fritsch, D.; Haala, N. SURE: Photogrammetric Surface Reconstruction from Imagery. In Proceedings of LC3D Workshop, Berlin, Germany, 4–5 December 2012.
34. Remondino, F.; Del Pizzo, S.; Kersten, T.; Troisi, S. Low-cost and Open-source Solutions for Automated Image Orientation—A Critical Overview. In Proceedings of Euromed, LNCS 7616, Lemesos, Cyprus, 29 October–3 November 2012; pp. 40–54.
35. Lingua, A.; Marenchino, D.; Nex, F. Performance analysis of the sift operator for automatic feature extraction and matching in photogrammetric applications. *Sensors* **2009**, *9*, 3745–3766.
36. Wu, C.; Agarwal, S.; Curless, B.; Seitz, S.M. Multicore Bundle Adjustment. In Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 21–23 June 2011; pp. 3057–3064.
37. Nocedal, J.; Wright, S. *Numerical Optimization*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2006.
38. Kazhdan, M.; Bolitho, M.; Hoppe, H. Poisson Surface Reconstruction. In Proceedings of the Fourth Eurographics Symposium on Geometry Processing, Sardinia, Italy, 26–28 June 2006.
39. Horn, B.K. Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am. A* **1987**, *4*, 629–642.